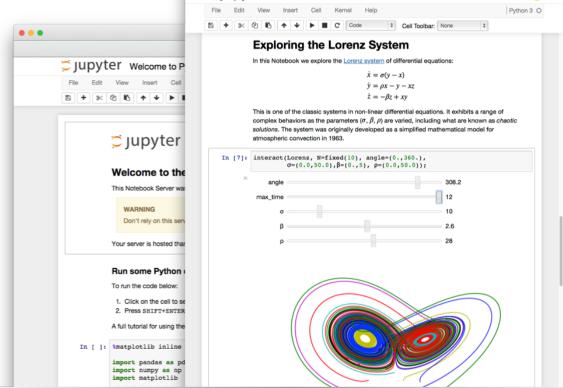
Sandboxed User Environments with Jupyterhub and Docker

Shreyas Cholia Donald Winston Daniel Gunter Lawrence Berkeley National Laboratory *Gateways 2016, San Diego*

Jupyter

- Interactive notebook environment that can combine code with visualizations, documentation and programmable widgets
- Formerly iPython Notebooks



CIUDVTET Lorenz Differential Equations (autosaved

2

Files Running Clusters

Select items to perform actions on them.

But Wait .. There's More

- Jupyter also includes
 - File Browser
 - Text Editor (with full syntax highlighting)
 - Terminal

| CJ | upyter potcar_setup.py - 09/05/2014 |
|------|--|
| File | e Edit View Language |
| | |
| 1 | #!/global/ul/s/shreyas/dev/testcif/bin/python2.7 |
| 2 | |
| 3 | import os |
| | import glob |
| 5 | import shutil |
| | import subprocess |
| 7 | |
| | pspdir = "" |
| 9 | |
| | count = 0 |
| 12 | <pre>while not os.path.exists(pspdir): if count != 0:</pre> |
| 13 | <pre>print("Invalid vasp dir!")</pre> |
| 14 | pspdir = raw input("Please enter full path where the POT GGA PAW PBE, " |
| 15 | "etc. subdirs are present. If you obtained the PSPs " |
| 16 | "directly from VASP, this should typically be the " |
| 17 | "directory that you untar the files to : ") |
| 18 | print |
| 19 | - |
| 20 | targetdir = raw_input("Please enter the fullpath of the where you want to " |
| 21 | "create your pymatgen resources directory: ") |
| | print |
| 23 | |
| | os.makedirs(targetdir) |
| | <pre>print("Generating pymatgen resources directory")</pre> |
| 26 | |
| 27 | <pre>for (parent, subdirs, files) in os.walk(pspdir): for subdir in subdirs:</pre> |
| 28 | <pre>for subdir in subdirs: filenames = glob.glob(os.path.join(parent, subdir, "POTCAR*"))</pre> |
| 30 | if len(filenames) > 0: |
| 31 | <pre>basedir = os.path.join(targetdir, os.path.basename(parent))</pre> |
| 32 | if not os.path.exists(basedir): |
| 33 | os.makedirs(basedir) |
| 34 | fname = filenames[0] |
| 35 | <pre>dest = os.path.join(basedir, os.path.basename(fname))</pre> |
| | |
| | |

| 👻 希 / global / homes / s / shreyas / dev |
|---|
| ۵. |
| D bdcclient-2.2.1 |
| C bdctest |
| C composetest |
| C data_docker |
| C db_scripts |
| C gridmap-verify |
| 🗅 Idap |
| 🗅 osg |
| 🗅 osg-pki-tools.orig |
| C testcif |
| C tmpvenv |
| 🛃 Untitled.ipynb |
| C osg-pki-tools-1.2.11-1.osg32.el5.noarch.rpm |
| 🗅 osg-pki-tools.tar.gz |

📁 jupyter

| total 16416 | | | | | | | | |
|-------------|----|---------|---------|----------|-----|----|-------|-----------------------|
| drwxrwxr-x | 2 | shreyas | shreyas | 512 | Jan | 29 | 2016 | build |
| drwxr-xr-x | 2 | shreyas | shreyas | 512 | Oct | 7 | 19:13 | conf |
| drwxr-xr-x | 3 | shreyas | shreyas | 4096 | Oct | 7 | 19:14 | conf.d |
| drwxrwxr-x | 4 | shreyas | shreyas | 4096 | May | 9 | 2013 | dist.eugridpma.info |
| drwxrwxr-x | 5 | shreyas | shreyas | 512 | Oct | 21 | 16:51 | foo |
| drwxrwxr-x | 5 | shreyas | shreyas | 512 | Oct | 21 | 23:01 | foobar |
| -rw | 1 | shreyas | shreyas | 16384 | Apr | 30 | 2015 | globus-user-env.sh.sw |
| drwxrwxr-x | 5 | shreyas | shreyas | 512 | Oct | 14 | 22:30 | msq |
| drwxrwxr-x | 5 | shreyas | shreyas | 512 | Jun | 3 | 22:25 | pmg |
| drwxrwxr-x | 6 | shreyas | shreyas | | | | | pydap-3.1.1 |
| drwxrwxr-x | 5 | shreyas | shreyas | 512 | Jan | 29 | 2016 | pydap-3.2 |
| drwxr-x | 18 | shreyas | shreyas | 4096 | May | 26 | 2015 | Python-2.7.10 |
| -rw-r | 1 | shreyas | shreyas | 16768806 | May | 23 | 2015 | Python-2.7.10.tgz |
| drwxr-x | 5 | shreyas | shrevas | 512 | May | 26 | 2015 | req |

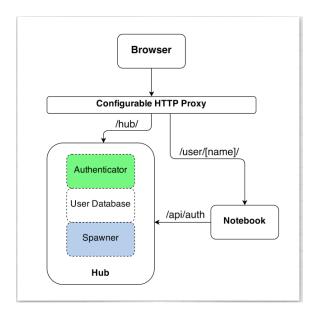
Self Contained environment

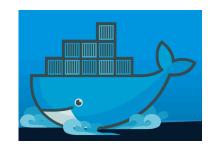
- Jupyter can essentially be a fully selfcontained interface for computation directly accessible through a web browser
- Fernando Perez has an exciting keynote tomorrow on the possibilities of Jupyter



Complementary Technologies

- Jupyterhub
 - Multiuser Jupyter service for launching notebooks
- Docker
 - Linux containers for creating private user sandboxed "OS" environments
 - You see your own private version of OS, filesystem, PIDs, users etc. i.e. virtualization-lite





Motivating Use Cases

- Materials Project: Jupyter for Workshops
- Containers with the entire materials project software stack pre-installed.
- Tutorial Notebooks
- Local Services (MongoDB and web services)



- IDAES: Jupyterhub for Persistent Sandboxes
- Capture complex software environment in a container
- Enable a development playground with worked examples for hacking on
- Serves as "live documentation"



DAFS

Goal

 Create a service that allows users to login and deploy their own private, pre-installed Jupyter notebook environments using containers.

Demo Time

We're going to

- Set up Jupyterhub
- Enable user authentication with Github Oauth
- Deploy notebooks in a per user Docker sandbox
- Run additional services in the sandbox
- Live notebook examples

Demo Links

If you want to try this at home

• Jupyterhub setup:

https://github.com/materialsproject/mp-jupyterhub/

• Docker Image:

https://hub.docker.com/r/materialsproject/jupyterhub-singleuser/

• Docker image build:

https://github.com/materialsproject/mp-jupyter-docker/

We're Hiring

- Postdoc to work on Jupyter for human-in-theloop HPC workflows
- Various CSE positions
- http://jobs.lbl.gov

Contact: scholia@lbl.gov